MONTHLY WEATHER REVIEW

Editor, James E. Caskey, Jr.

VOL. 77, No. 9
W. B. No. 1545

NOVEMBER 1949

CLOSED JANUARY 5, 1949
ISSUED FEBRUARY 15, 1949

# PERSISTENCY OF RAIN AND NO-RAIN PERIODS DURING THE WINTER AT SAN FRANCISCO

By DONALD L. JORGENSEN

[Weather Bureau Airport Station, San Bruno, Calif.]

Manuscript received July 8, 1949

ABSTRACT

A 20-year record of winter precipitation data for San Francisco is used to determine frequencies of periods of rain and no-rain of various lengths and the frequencies of periods of a given type of weather followed by the same type of weather. The observed frequencies are compared with those expected on chance and are examined for evidence of persistence of rain and no-rain. The use of persistency as a forecasting aid is discussed and its skill determined.

## INTRODUCTION

The Northeast Pacific with its great expanse of water surface has been found to be especially well suited for the development of forecasting methods based on weather map types. The relatively uniform north-south temperature gradient of the ocean surface and the lack of orographic influences do not lead to abrupt changes in the circulation patterns. Air masses moving into the area from the continental source regions adjacent to the Pacific Ocean have undergone considerable modification before becoming involved in the existing major air streams. The same considerations which have led to successful weather map typing for the area lend themselves to a study of persistence. Accordingly, it was felt that a study of persistency of weather at San Francisco held promise of worthwhile results.

## PROBLEM

The weather along the Pacific coast in the vicinity of San Francisco is controlled to a large extent by the position, shape, and intensity of the Eastern Pacific subtropical High, and to a lesser extent by the location and intensity of the Aleutian Low. Forecasting studies for the west coast have recognized the Pacific High as the predominant controlling feature of the surface map. Vernon [1] in his study of winter precipitation at San Francisco based on map types found that the most important considerations were the position of highest pressure off the California coast in determining the probability of rain during a Westerly Type and the shape of the high and its east-west displacement leading to the Northerly or Southerly Type classification. In the Westerly Type the High acts usually as a barrier to storms approaching the central Pacific coast, but if the center of the High is sufficiently far south such storms may approach the coast in the vicinity of San Francisco. In the Northerly Type, the High acts as a steering mechanism to bring Lows forming in the Gulf of Alaska southward along the Pacific coast; in the Southerly Type, storms forming in the Central Pacific are carried northward. During the relatively infrequent instances of a storm approaching from the east, the Pacific High again acts in its steering capacity.

Forecasters have long recognized the Eastern Pacific High as a persistent phenomenon. While acting as a control to the surface features, it is in turn probably controlled by air streams prevailing at middle and upper levels in the atmosphere. If the forecaster accepts this viewpoint, he is faced with the problem, except for variations of the pressure distribution as a result of migrating ridges, of forecasting surface changes which are due to causes only partially known and of such nature that data describing them are unavailable. When confronted with this situation, the forecaster generally tends to fall back on the known persistency of the main features of the weather map and to continue forecasting the current weather trend. With the exception of the study of the persistency of weather at several European stations by Newnham in 1916 [2] and Besson in 1924 [3], and at a station in the central United States by Blair in 1924 [4], very little information is available in a satisfactory form to enable the forecaster to use day-to-day persistency in other than a subjective manner. Blumenstock [5], in his analysis of drought by means of the theory of probability, has made a general study of the probability of occurrence of periods with little or no precipitation over most of the United States, but the study does not lend itself for use by the forecaster. The recent emphasis placed on the objectivity of forecasting methods brings to light the importance of determining (1) the degree to which persistency resulting from broadscale flow patterns enters into the forecasting problem, and (2) the way in which it contributes to the indicated skill of an objective method. The purpose of this paper is to present the results of a study of these two aspects of persistency of rain and no-rain periods at San Francisco.

## PROCEDURE

For the purposes of the study persistency was defined as follows: Persistency is the observed percentage frequency in excess of chance that a period of $M$ days of a given type of weather is followed by $N$ or more consecutive days of the same type of weather beginning $L$ days later. (In this definition a period is taken as beginning with the last change in the weather type.)

The period chosen for the study includes the months December through March for the 20 years 1927–47. The daily weather record for San Francisco published in the "Climatological Data" for California was used to type each day as either a rain or a no-rain day. A day receiving a measurable amount of rain was considered to be a rain day. As an initial step in the study, the frequencies of periods of rain and no-rain of various lengths were determined. Next, from these data, the frequencies of periods of various lengths of rain and no-rain followed by one or more days, two or more days, etc., of the same type of weather were determined. Through a further consideration of the record, frequencies of recurrence of a given type of weather on the first, second, and third days following periods of various lengths were obtained.

During the 80 months (December through March) of the 20-year record considered, 819 rain days occurred out of a total of 2,425 days, thus giving 33.8 percent rain days. For determining the probability on chance of rain occurring on any one day a ratio of $\frac{1}{3}$ was used. On this basis, the probability of a rain period continuing unbroken for 1 day is $\frac{1}{3}$, for 2 days $\frac{1}{9}$, for 3 days $\frac{1}{27}$, etc. Similarly, the probability on chance that a no-rain period will continue for one more day is $\frac{2}{3}$, for two more days $\frac{4}{9}$, for three more days $\frac{8}{27}$, etc.

In counting the periods of rain and no-rain during the winter season, the periods at the beginning and end of the season were arbitrarily assigned to that month in which the greatest number of days fell. For this reason the total number of days involved in a portion of the study differs slightly from the total number recorded during the months of the 20 years considered.

## FREQUENCY OF PERIODS OF RAIN AND NO-RAIN

The results obtained from the tabulation of the frequencies of rain periods of various lengths are shown in table 1. Also shown in this table are the frequencies of the periods of various lengths expected on a chance distribution of the rain days and the cumulative totals of the observed rain periods from the longest to the shortest.

The calculation of the frequency of periods of various lengths expected on chance requires the solution of the problem of runs for two types of weather. Cochran [6] gives the equation

$$f_{r,m}=p^r q[2+q(m-r-1)] \qquad (1)$$

where $f_{r,m}$ is the frequency of runs of length $r$ out of $m$ trials, for the type of weather having $p$ as the probability of occurrence during the unit period, $q=1-p$ and $1\leq r\leq(m-1)$. This equation was used in obtaining the frequencies of rain periods of various lengths with $p=\frac{1}{3}$, $q=\frac{2}{3}$, and $m=121.3$, the average number of days per season. Since 20 years of record were used, the frequencies thus obtained were multiplied by 20.

The observed total number of rain periods of different lengths (table 1) indicates an apparent tendency for rain to persist for about 5 days, as shown by the fact that the number of periods lasting 5 days actually exceeded the number lasting 4 days. The occurrence of periods lasting 1 and 2 days was less frequent than that expected on chance, and the number lasting longer than 3 days was in general more frequent than expected on chance. The average number of rain periods per month was 3.84. In the 307 rain periods considered during the 20 years of record, 810 days of rain were involved giving an average length for the rain periods of 2.64 days. The longest rain period was 19 days.

Data for the no-rain periods are given in table 2. Equation (1) with $p$ and $q$ interchanged was used to calculate the number of runs of various lengths expected on chance. The frequency of observed no-rain periods increases from 7 periods lasting 8 and 9 days to 12 periods lasting 10 days and 11 periods lasting 11 days and thus indicates an apparent tendency for the no-rain periods to persist about 10 days. The average number of no-rain periods per month was 3.89. The small difference compared to the rain periods is due to the fact that there tended to be a higher ratio of no-rain to rain periods at the beginnings and ends of the seasons. In the 311 no-rain periods considered, 1,618 no-rain days were involved, giving 5.20 days as the average length of the no-rain periods. The periods lasting from 1 to 9 days were considerably less frequent than would be expected on chance,

TABLE 1.—*Observed frequency of rain periods by duration in days, frequency of rain periods of various lengths expected on chance, and frequency of rain periods cumulated from the longest to shortest (20 seasons, 1927–47)*

| | Length of rain periods in days | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 |
| December | 33 | 15 | 9 | 5 | 8 | 2 | ---- | ---- | ---- | 2 | ---- | 1 | ---- | 1 | ---- | ---- | ---- | ---- | ---- |
| January | 28 | 16 | 14 | 3 | 7 | 5 | 1 | 3 | ---- | ---- | ---- | ---- | ---- | ---- | ---- | ---- | ---- | ---- | ---- |
| February | 34 | 25 | 8 | 4 | 3 | 2 | 2 | 2 | ---- | ---- | ---- | ---- | ---- | ---- | 1 | ---- | ---- | ---- | 1 |
| March | 33 | 14 | 7 | 8 | 3 | 3 | 2 | 2 | ---- | ---- | ---- | ---- | ---- | ---- | ---- | ---- | ---- | ---- | ---- |
| Observed total | 128 | 70 | 38 | 20 | 21 | 12 | 5 | 7 | 0 | 2 | 0 | 1 | 0 | 1 | 1 | 0 | 0 | 0 | 1 |
| Frequency expected on chance | 362.4 | 119.8 | 39.6 | 13.1 | 4.3 | 1.4 | 0.47 | 0.16 | 0.05 | 0.02 | 0.01 | <.01 | <.01 | ---- | ---- | ---- | ---- | ---- | ---- |
| Observed cumulative total (reverse) | 307 | 179 | 109 | 71 | 51 | 30 | 18 | 13 | 6 | 6 | 4 | 4 | 3 | 3 | 2 | 1 | 1 | 1 | 1 |

TABLE 2.—*Observed frequency of no-rain periods by duration in days, frequency of no-rain periods of various lengths expected on chance, and frequency of no-rain periods cumulated from the longest to shortest (20 seasons, 1927–47)*

| | Length of no-rain periods in days | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 | 22 | 23 | 24 | 25 | 26 | 27 | 28 | 29 | 30 |
| December | 27 | 11 | 8 | 6 | ---- | 2 | 5 | 2 | 2 | 5 | 1 | 4 | 1 | ---- | 3 | 1 | 1 | 1 | ---- | ---- | ---- | ---- | ---- | ---- | ---- | ---- | 1 | 1 | ---- | ---- |
| January | 22 | 11 | 10 | 6 | 3 | 3 | 4 | 3 | 1 | 3 | 5 | 2 | 1 | ---- | 1 | ---- | ---- | ---- | ---- | 1 | 1 | ---- | ---- | ---- | 1 | ---- | ---- | ---- | ---- | ---- |
| February | 19 | 16 | 12 | 9 | 3 | 4 | 2 | 1 | 2 | 1 | 1 | ---- | 2 | ---- | ---- | 1 | 1 | ---- | 1 | 1 | 1 | 1 | ---- | ---- | ---- | ---- | ---- | ---- | ---- | ---- |
| March | 18 | 13 | 7 | 6 | 7 | 2 | 3 | 1 | 2 | 3 | 4 | 1 | 1 | ---- | ---- | 1 | ---- | ---- | ---- | ---- | ---- | ---- | 1 | ---- | ---- | ---- | ---- | 1 | 1 | 1 |
| Observed total | 86 | 51 | 37 | 27 | 13 | 11 | 14 | 7 | 7 | 12 | 11 | 7 | 5 | 0 | 4 | 3 | 2 | 1 | 1 | 2 | 2 | 1 | 1 | 0 | 1 | 0 | 1 | 2 | 1 | 1 |
| Frequency expected on chance | 185.6 | 122.8 | 81.2 | 53.7 | 35.5 | 23.5 | 15.5 | 10.3 | 6.8 | 4.5 | 3.0 | 2.0 | 1.3 | .85 | .56 | .37 | .25 | .16 | .11 | .07 | .05 | .03 | .02 | .01 | <.01 | <.01 | ---- | ---- | ---- | ---- |
| Observed cumulative total (reverse) | 311 | 225 | 174 | 137 | 110 | 97 | 86 | 72 | 65 | 58 | 46 | 35 | 28 | 23 | 23 | 19 | 16 | 14 | 13 | 12 | 10 | 8 | 7 | 6 | 6 | 5 | 5 | 4 | 2 | 1 |

and the periods longer than 9 days were on an average much more frequent. The longest period without rain recorded was 30 days.

A further consideration of equation (1) shows that for commonly observed values of $p$ and $q$, and $m \gg r+1$, the ratio of the no-rain to the rain periods expected on chance approaches the geometric progression:

$$\frac{\text{frequency of no-rain periods of length } r}{\text{frequency of rain periods of length } r}=\left(\frac{q}{p}\right)^{r-2} \qquad (2)$$

As previously indicated for this study, $p=⅓$, and $q=⅔$; thus,

$$\left(\frac{q}{p}\right)^{r-2}=(2)^{r-2}$$

A comparison of the ratios expected on chance and those actually observed is shown in table 3. The ratios expected on chance increase much more rapidly than the observed ratios. Since the exponent of equation (2) is zero for $r=2$, this indicates that on a chance basis the frequency of periods of no-rain lasting 2 days is the same as the frequency of rain periods of the same length regardless of the climatological expectancy (within, of course, the limits of $p$ and $q$ for which the approximate relation (2) is valid). A comparison of the observed and expected ratios for this value of $r$ shows that this condition is not attained by the actual data.

TABLE 3.—*Ratios of frequency of no-rain periods to frequency of rain periods expected on chance compared with observed ratios*

| Days in period ($r$) | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Ratios expected on chance | 0.5 | 1 | 2 | 4 | 8 | 16 | 32 | 64 | 128 | 256 | 512 |
| Observed ratios | 0.67 | 0.73 | 0.97 | 1.4 | 0.62 | 0.92 | 2.8 | 1.0 | ---- | (6.0) | ---- |

## FREQUENCY OF RAIN AND NO-RAIN PERIODS FOLLOWED BY THE SAME TYPE OF WEATHER

From the accumulated values obtained from tables 1 and 2, the percentage of times a period of a given type of weather (rain or no-rain) is followed by the same type of weather may be calculated. From table 1, for example, out of the 307 cases of a rain period of at least 1 day duration, 179 cases, or 58 percent, were of at least 2 days duration; 109 cases, or 36 percent, were of at least 3 days duration; etc. Of the 179 cases lasting 2 days or more, 109, or 61 percent, lasted at least 3 days; 71 cases, or 40 percent, lasted at least 4 days; etc. Corresponding frequencies were obtained in a similar manner for the no-rain periods. Data obtained in this manner were plotted for rain periods and no-rain periods, and lines of equal percentage frequencies were drawn (figs. 1 and 2). Percentage frequencies of a rain period or no-rain period continuing for 1, 2, . . . days on a chance basis were entered at the tops of the respective graphs.

Figures 1 and 2 show that the probability that existing weather at San Francisco will recur is considerably greater than would be expected by chance. The greatest probability of a rain period continuing for at least one more day is attained following a rain period of 4 days duration with a value of 72 percent as compared to 33 percent expected on chance. The greatest probability of rain continuing for at least two more days occurs following a rain period of 3 days. The probability here is 47 percent compared to 11 percent expected on chance. Figure 2

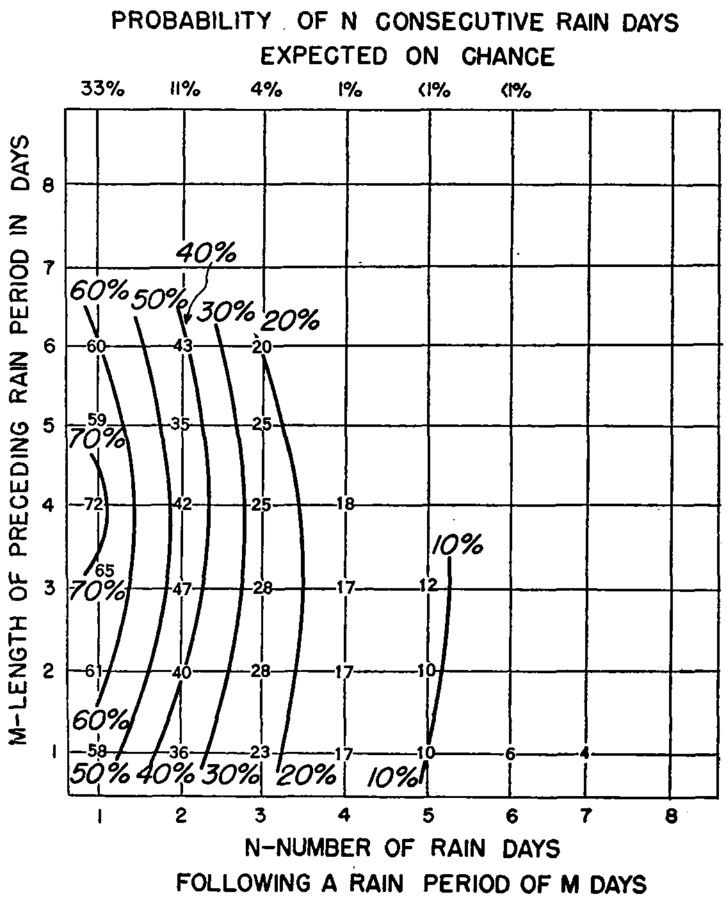


FIGURE 1.—Graph showing isograms of percentage frequency of occurrence of a rain period of $N$ or more days following a rain period of $M$ days. Numbers at top of graph give the probability of $N$ consecutive rain days expected on chance.

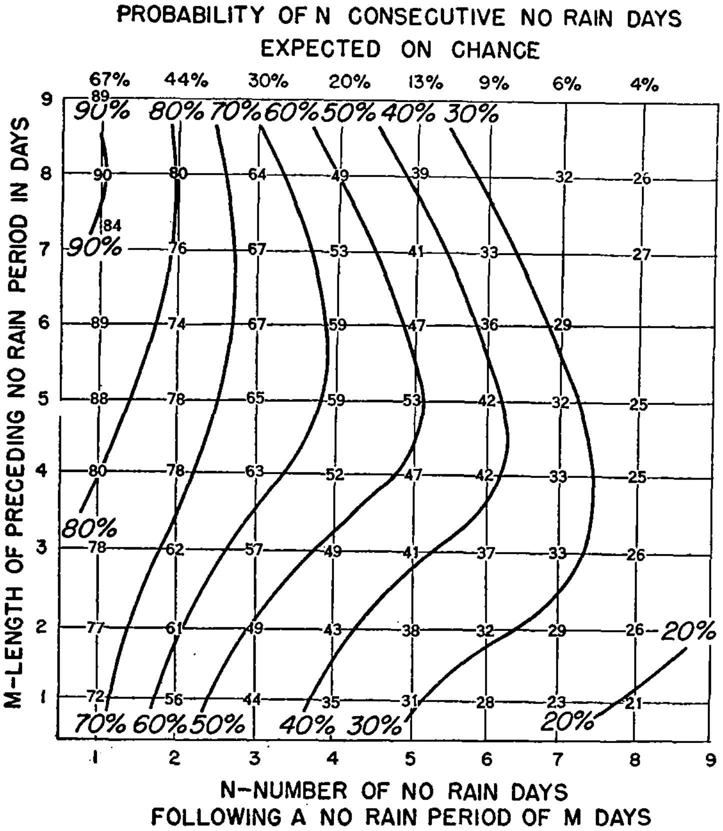


FIGURE 2.—Graph showing isograms of percentage frequency of occurrence of no-rain period of $N$ or more days following a no-rain period of $M$ days. Numbers at top of graph give the probability of $N$ consecutive no-rain days expected on chance.

shows that the greatest probability of a no-rain period continuing for at least one more day is 90 percent attained following a no-rain period lasting 8 days; the probability expected on chance is 67 percent. Other combinations may be read from the two charts. Data for rain periods lasting longer than 8 days and no-rain periods lasting longer than about 13 days were insufficient to give satisfactorily consistent values for the frequencies.

## PERSISTENCY AS A FORECASTING AID

Persistency, as defined in this paper, gives a measure of the tendency *in excess of chance* for weather to continue unchanged, i.e., the probability of continuance obtained from the observed record minus the probability calculated for chance on the observed ratio of occurrence of the weather types. On this basis, persistency of rain may be obtained from figure 1 and no-rain from figure 2 for various values of $M$ and $N$, with $L$ equal to 1 (where $M$, $N$, and $L$ are used as given in the definition of persistency, p. 304).

In actual practice it is necessary to forecast for the second or third day in the future without having definite knowledge of the weather on the intervening days. In other words $L$ in the definition takes on values greater than 1. In figure 3 persistency obtained from figure 1 for $L$ and $N$ equal to 1 was plotted to give the curve labeled $L=1$. Curves labeled $L=2$ and $L=3$ were drawn from data obtained from a further examination of the record (data not tabulated). This figure shows that persistency reaches a maximum on the fifth day of a rain period with a value of 39 percent. Expressed in a different way, if "today" is the fourth day of a rain period the over-all probability that it will rain "tomorrow" is 72 percent (from fig. 1), which, when reduced by the 33 percent

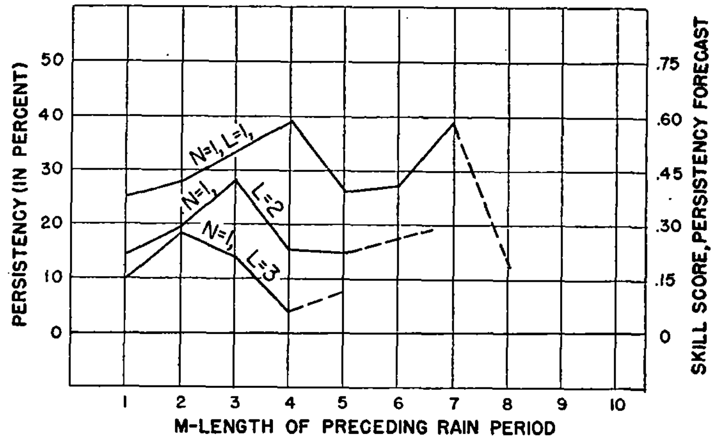


FIGURE 3.—Graph showing persistency of rain on first, second, and third days ($L=1, 2, 3$) following a rain period of $M$ days and corresponding skill score of a persistency forecast.

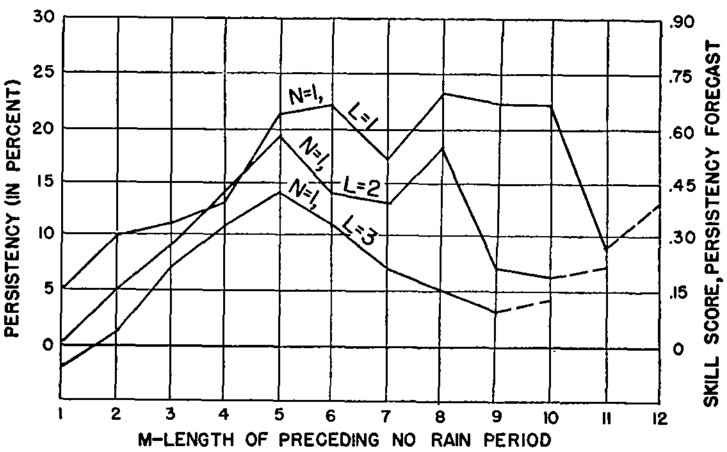


FIGURE 4.—Graph showing persistency of no-rain on the first, second, and third days ($L=1, 2, 3$) following a no-rain period of $M$ days and corresponding skill score of a persistency forecast.

probability that it will rain "tomorrow" due to chance, gives 39 percent for the value of persistency. Under the same conditions, i. e., rain during the last 4 days, the probability that it will rain "the day after tomorrow," without considering the weather "tomorrow," is 48 percent as obtained from a study of the record. Reducing this figure again by 33 percent, the probability of rain due to chance, gives 15 percent, which is the value read from the curve labeled $L=2$ in the figure. Persistency existing for other situations may be taken from the figure.

Corresponding curves for the no-rain periods are given in figure 4. It is of interest to note the small values of persistency (in one case actually negative) which occur for the no-rain periods for small values of $M$. This fact may be interpreted to mean that a period of one or two no-rain days is not sufficient to establish a significant no-rain trend. It is not until a no-rain period has continued for about 6 days that the persistency begins to approach the highest value; the maximum persistency appears to occur on the ninth day.

It is of interest to compute the skill of "persistency forecasts." With the development of objective forecasting methods, the need for a measure of the success of a method has led to the development and use of a specialized type of skill score. This score is designed to measure the residual skill by the removal from consideration of those successful forecasts which would have been successful on a chance basis even though the forecaster had no meteorological information available. Thus, the forecaster is not given credit for the number of hits which would be attained with a purely random distribution of his forecasts throughout the month. On this basis for the determination of skill, a persistency forecast will attain a worthwhile score. The skill score $S_c$ as generally used may be defined as follows: $S_c=\frac{C-E_c}{T-E_c}$, where $C$ is the number of correct forecasts; $E_c$ the number of forecasts expected to be correct due to chance; and $T$ the total number of forecasts. Out of the 100 persistency forecasts, the percentage frequencies obtained from figures 1 and 2 give the number of correct forecasts, $C$, for those situations to which the charts apply. The number of correct forecasts out of 100 expected to be correct due to chance, $E_c$, is given by the percentage at the tops of the graphs. By our definition, persistency is the difference between these two percentages. Expressing persistency, $P$, as a decimal we may write the equality $PT=C-E_c$. Hence, $S_c=\frac{PT}{T-E_c}$. For $N=1$, the most common type of forecast, $E_c=T/3$ for the rain cases at San Francisco, and for the no-rain cases $E_c=2T/3$. Therefore, for the persistency rain forecasts, $S_c=3P/2$, where $P$ is expressed as a decimal; and for the persistency no-rain forecasts, similarly, $S_c=3P$.

It is seen therefore that, for a given value of $N$, the skill score may be obtained by multiplying the value of persistency expressed as a decimal by a constant factor. Scales giving the skill scores of persistency forecasts corresponding to the persistencies indicated have been entered at the right-hand edges of figures 3 and 4. These figures indicate that for the number of days considered, the skill of persistency rain forecasts varies from about .10 up to .55, and for persistency no-rain forecasts from slightly negative values to about .70.

An average skill score for day-to-day persistency forecasts may be obtained through a consideration of the average number of rain periods per month. For a 30-day month this number is about 3.8. Assuming a missed forecast at the beginning and end of each period, the total misses per month would be 7.6. Calculating the skill score on this basis gives:

$$S_c=\frac{(30-7.6)-(\frac{1}{3}\times 10+\frac{2}{3}\times 20)}{30-(\frac{1}{3}\times 10+\frac{2}{3}\times 20)}=.43$$

## CONCLUSIONS

1. Persistency is a very real meteorological phenomenon and deserves greater attention in the development of objective forecasting methods.

2. The importance of persistency probably rests on the fact that there are obscure factors involved in weather situations which are beyond the scope of expression by common meteorological variables. Persistency gives a rough integrated value of all variables. In this respect, persistency may be considered as a meteorological variable.

3. The study of persistency at a station provides the forecaster with an objective presentation which will aid him in evaluating the current weather trends.

4. A knowledge of persistency is essential to evaluate properly the skill of a forecasting method as evidenced by the skill score. For example, the difference between the attained skill score and the score due to persistency gives an indication of the ability of the method to forecast changes. If this difference is negative the method is not showing worthwhile skill.

## REFERENCES


1. E. M. Vernon, "An Objective Method of Forecasting Precipitation 24–48 Hours in Advance at San Francisco, Calif.," *Monthly Weather Review*, vol. 75, No. 11, November 1947, pp. 211–219.
2. E. V. Newnham, "The Persistence of Wet and Dry Weather," *Quarterly Journal of the Royal Meteorological Society*, vol. 42, No. 179, July 1916, pp. 153–162.
3. L. Besson, "On the Probability of Rain," *Monthly Weather Review*, vol. 52, No. 6, June 1924, p. 308.
4. T. A. Blair, "The Coefficient of Persistence," *Monthly Weather Review*, vol. 52, No. 7, July 1924, p. 350.
5. G. Blumenstock, Jr., "Drought in the United States Analyzed by Means of the Theory of Probability," *Technical Bulletin No. 819*, U. S. Department of Agriculture, April 1942.
6. W. G. Cochran, "An Extension of Gold's Method of Examining the Apparent Persistence of One Type of Weather," *Quarterly Journal of the Royal Meteorological Society*, vol. 64, No. 277, October 1938, pp. 631–634.